

論文内容の要旨

博士論文題目

Bottom-Up Multi-Agent Reinforcement Learning for Complexity and Uncertainty

(複雑性と不確実性に対処するボトムアップマルチエージェント強化学習)

氏名

青谷 拓海

(論文内容の要旨)

マルチエージェントシステムは一台のエージェントでは成し遂げられないタスクを複数台が協調することで解決する可能性を秘めており、様々な実世界問題への適用が期待されている。しかし、そのシステムの複雑さと不確かさはエージェント数に伴い爆発的に増加するため、システム全体に対する制御戦略を事前設計することは難しい。マルチエージェント強化学習 (MARL) は経験データより各エージェントの制御戦略を学習・獲得するため、この問題に対処できるはずだが、実際には、既存の MARL は対象とするシステムを共通タスクを持ち中央集権システムを内包したものに限定してしまっている。

そこで、本論文では各エージェントが個別タスクを有する自律分散システムに対しても適用可能な「ボトムアップ MARL」を提唱する。この枠組みであれば従来の MARL より対象システムを大幅に拡大することが可能となる一方で、複雑性と不確実性の問題はより顕著になり集団行動の獲得や安全な学習を阻害してしまう。これらに対処すべく、本研究では以下の 2 つの新たな学習アルゴリズムを開発する。

- i) 集団行動を獲得するための 4 つのコンポーネントで構成される報酬整形アルゴリズムを開発する。各エージェントは個別の報酬関数を保有しているが、そこから得られる報酬値を学習時には他エージェントと共有することを仮定する。この共有された情報をもとに、他エージェントの報酬およびその統計量 (分散・相関) を予測可能とする。また、予測精度を高めるための学習促進項を合わせて設計する。これらを組み合わせ、

正の相関が高く報酬の分散も小さい，すなわち協力可能なエージェントの報酬を重視して足し合わせた値を新たな報酬として与える．この総和の最大化により互いに協力可能な範囲で集団行動を学習する．このアルゴリズムの有効性をシミュレーションおよび実機実験により検証する．

- ii) 不確実性の高い環境下での長期予測により安全性を保証するためのモデル学習アルゴリズムを開発する．長期予測では，予測区間中に一度でも大きな予測誤差が生じると，その後の予測が破綻してしまう一方で，各予測の予測誤差の蓄積も抑制しなければならない．すなわち，予測誤差に関する偏差と分散のトレードオフを適切に考慮しながらモデルを学習することを目指す．まず，このトレードオフを陽に扱える新しい最小化問題を拡大チェビシェフ法を用いて導入する．その上で，偏差と分散のトレードオフを司るハイパーパラメータをモデル学習と並行してメタ最適化すべく，方策勾配法を活用したアルゴリズムを設計する．このアルゴリズムの有効性をシミュレーションにより検証する．

氏名	青谷 拓海
----	-------

(論文審査結果の要旨)

従来のマルチエージェント強化学習では、一般的には全エージェントで共通のタスク（報酬関数）が与えられているため、集団行動の獲得は比較的容易な一方で、全エージェントの情報に基づいた報酬関数はいわば中央集権システムとなってしまう。このため扱えるエージェント数には限りがある。本論文ではこの根源的な課題を新たな枠組みであるボトムアップマルチエージェント強化学習を導入することで回避している。この場合、各エージェントが自身の情報のみに基づいた報酬関数のみを保有するため、システムは自律分散性を得る。一方で、集団としてのタスクが陽に与えられないため、他エージェントとの協調が生まれず、また、自律分散システムに起因する部分観測問題により不確実性が高くなってしまう。これら 2 つの課題に対して、本論文ではそれぞれに対応したアルゴリズムを開発している。

1 つめの報酬整形アルゴリズムは、最小限のコミュニケーションから他エージェントとの関係性を数値的に類推し集団行動を創発させる実用性の高いものであり、基盤となる強化学習アルゴリズムに依らない一般性の高さも評価できる。また、シミュレーションだけでなく実機実験でも検証を実施しており、十分な価値のある結果とみなせる。

2 つめのモデル学習アルゴリズムでは、学習における一般的な課題とされる偏差と分散のトレードオフに上手く着目した上で、それを陽に扱える最適化問題の定式化や最適なトレードオフのバランスを見出すメタ最適化手法の開発と、堅実な問題解決を達成している。検証はシミュレーションに留まっているものの、マルチエージェントシステムに限らず人が介在するシステムを扱うなど、一般性の高さを示す結果となっており、高く評価できる。

このように、本論文はマルチエージェント強化学習における新しい枠組みの提案、およびそこで生じる課題を汎用的なアルゴリズムの開発により解決していることに鑑み、新規性および有用性の観点から一定の学術的意義があるものと評価できる。よって、本論文は博士（工学）の学位論文として価値あるものと認める。